Comprehensive Examination of Miss Pregya Poonia, Ph.D(computer Sc.) Student

Research Proposal on "**Short Term Traffic Flow Prediction Using Deep Learning Approach**"

MODY UNIVERSITY
OF SCIENCE AND TECHNOLOGY
LAKSHMANGARH • RAJASTHAN
A Leading Women's University

Guided by

Prof. V. K. Jain (DEAN CET)

Prof. Anil Kumar

Short Term Traffic Flow Prediction Using Deep Learning Approach

1. INTRODUCTION

Cumulating high quality traffic, data information is the key factor to attain the performance of Intelligent Transport System (ITS). Prefiguring the future patterns in traffic flow turns into highly vital in Advanced Traffic Management (ATM) and advanced Traveller Information System (ATIS). Currently, prefiguring the traffic speed is the main fundamental study in the area of traffic research community. By using the forecasted data, such as travel time data, traffic condition entropy and traffic volume data, travellers can replan their travelling ways to save the time and cost. Basically, predicting the speed of traffic is a part of the traffic information prediction. Based on the current and past traffic observations, the traffic speed can be predicted using two-terms:

Short-term traffic flow forecasting, which involves the forecasting of traffic mass in the following time interval generally in the compass of five minutes to one hour. And long-term traffic flow prefigure is founded on hours, days, and months and even on years.

Short-term flow prediction is the paramount research problem in the sphere of TMS. Many researchers are showing their efforts in this field from past two decades. The count of vehicles moving at a particular point per unit time period is a case of random process which comprises of a set of isolated points hoarded over time. Traffic engineering analyses the deportment of traffic and also design better adeptness for a safe, economical and smooth traffic operation. Presently, relevant researches primarily divided up into two different sections: parametric modelling and non-parametric modelling. As the flow of water, traffic flow also has several parameters allied within it.

The growth in traffic in developing countries is epidemic and will be continue to be so in coming years. Smart cities have deployed latest technologies to gratify the necessarily of travellers through effective navigation. As per to a latest survey approximately 60 percent of people will live in cities by 2050 and as a result, millions of vehicles (cars and two wheelers) will run on the roads. The basic feature of ITS is dynamism of phenomena and the dynamism is an aftermath of the transport system properties: high level of complexity, large number of

users, occurrence of complex and random traffic phenomena and susceptibility to external factors, these are the natural features of ITS.

2. Literature Review

2.1 Short-term traffic prediction

In smart cities it has become an important subject of research in transport engineering and the urban logistics system in the past decades. The main aim is to uprise intelligent traffic control methods, create advanced passenger information system and applications which facilitate the deliveries of goods in the municipal road network. Except this the other reason which we can consider is growing traffic in urban areas and the development of data acquisition and processing technologies.

As an integral part of ITS, Short-term traffic prefigure concludes future value of parameters such as traffic volume, density, and speed or estimated travel time in divisions of the urban road networks. Such prediction has the time horizon spanning which can go from a few seconds to maximum a few hours.

2.2Review of traffic flow prediction technologies

There are two methods involved in traffic flow foresee, these are Parametric and Non Parametric approaches. The famous approach is non-parametric and being highly used in current research studies. Many technologies are there including statistical models and machine learning approaches.

Among many parametric techniques, linear Auto-Regressive Integrated Moving Average (ARIMA) and some modified models have been applied in many research studies. At starting, the ARIMA was first launched in traffic information prediction to investigate the stochastic feature of traffic system. The subset ARIMA, ARIMA with explanatory variables and seasonal ARIMA were presented considering various features of the traffic flow, these are developed for the extension of ARIMA. Seasonal ARIMA is used in India in a recent research work by S. Vasantha Kumar (2015), Lelitha Vanajakshi (2015). [2]SARIMA model used only special input data to detect short term prediction of traffic flow. For this Chennai's

3-lane roadway was selected, three days data study was used for the model development using SARIMA. Later on, the autocorrelation function (ACF) and partial autocorrelation function (PACF) were plotted to identify the suitable order of SARIMA model. The proposed model was formalised by performing 24 hrs leading forecast and the prediction flows were equated with the actual flow values [3]. Short term traffic flow prediction was done on the basis of historical and actual time data during the peak periods of morning and evening times. The equated mean absolute percentage error (MAPE) of real and predicted flow showed in the compass of 4-10 which is admitted in most of the ITS applications. The highlighted issue with this model was database.

Various Machine Learning techniques are involved in traffic flow prediction and various Deep Learning models have been introduced in this field. A Deep Learning method, Deep Belief Network was proposed by Yuhan Jia et al. (2016) for short-term traffic speed information prediction. The DBN model is trained in a greedy unsupervised method and fine-tuned by labelled data. Data is gathered from one arterial in Beijing, China. DBN is comprised of a stack of RBM, which is considered as the family of belief networks. On the tip of RBMs, there is a BPNN regarded as the output layer, which can fine-tune the entire model by using back-propagation algorithm. Three forecasting time horizons are selected in this study, including 2-minute, 10-minute and 30-minute horizons. Three performance measurements are selected: (a) the Mean Absolute Percentage Error (MAPE), (b) the Root Mean Square Error (RMSE), and (c) the Normalized Root Mean Square Error (RMSE). Problem occurs when the prediction horizons are longer; more stochastic features are missing from the DBN results [5]. It shows poor performances during the evening's peak hours.

Han Jiang et al. (2016) examined the multistep ahead prediction functioning by using eight dissimilar models and took the collected data of 2-minute travel speed from three Remote Traffic Microwave Sensors which is situated on a southbound segment of 4th ring road in Beijing, China. This model was based on the theory of machine learning and statistical. Five machine learning models were considered: Back Propagation Neural Network (BPNN), Nonlinear Autoregressive Model with Exogenous Inputs Neural Network (NARXNN), Support Vector Machine with Redial Basic Function as kernel function (SVM-RBF), Support Vector Machine with Linear Function (SVM-LIN), and Multilinear Regression (MLR) as candidate. Three statistical models: Autoregressive Integrated Moving Average (ARIMA), Vector Autoregression (VAR), and Space-Time (ST) model are also selected. The following meaningful outcomes were found from the predicted results: (1) Speed deteriorates prediction

accuracy increases for all models, (2) two statistical models can clearly be outperformed by the BPNN, NARXNN and SVM-RBF, they are: ARIMA and VAR, (3) ANN has superior prediction performance rather than SVM and MLR, (4) the ST model can give the lowest MAE comparing with ARIMA and VAR. This case study supplies useful applications of speed prediction and comparisons of prediction performance.

Zhiheng Yu (2015) proposed two combinational prediction models which rely on GM, ARIMA and GRNN. GM (Grey Model) is, to extract important information via the partly known information processing and extending. This model was designed so that the accuracy of traffic flow prediction can be improved. For certain distribution GM does not require large samples on the data to be subjected. ARIMA is a method of a time series prefigure. It can predict the curve for short-term and long-term forecast and it has smaller prediction error. GRNN (generalized regression neural network) which is based on non-parametric regression has a nonlinear mapping power which solves the curve fitting issue. GRNN is able to distinguish the concealed rules from the large amount of data and then integrate this learning information into the linking weights between neurons, but it gives the poor prediction accuracy when the training samples are less. The outcome of recently developed combinational forecast model was analyzed by taking 54 pieces of main road traffic data on Chongzun of China. The grey system theory modelling software 3.0 and MATLAB R2012b are used in these experiments. This comparison showed either combinational model has mellow prediction accuracy than any other individual one. But still, how to opt the earmark model for combination is a question.

A current research paper by Daniel Kubek (2016) presented an approach to robust optimization which shows the logistics vehicle routes in urban areas on the basis of figured short-term traffic time forecasting in a particular selected area of the urban road network. Spectral analysis model had been used to specify the calculated values of optimization parameters. In urban network the robust routes for logistics vehicles are used to determine with the help of the shortest path problem. The uncertainty set is forged on the basis of prefigure travel times in chosen divisions, estimated by means of spectral analysis. The classic approach have been equated with the attained data wherein it is contemplated that the optimization parameters are certain and authentic. The quality of final results is affected with the function of forecasting proficiency with robust optimization models. Transport and logistical companies can take the advantage of this approach in their practical operation.

In the past few years many researchers are being attracted towards deep learning including academics and industries. Recently, deep learning neural network algorithm is used to formulate three models for traffic flow prediction. These models were designed by Sivabalaji Manoharam (2016), this was a research project. In this project a multi layered structure was used to extract meaningful data through a deep learning algorithm and the data used in this process was weather data along with traffic flow pattern. The processing data was taken from the United Kingdom website which is an opened data source data website. The traffic data consists of 23 loop and the weather data is obtained on one hour time interval and it's located near Leeds city. In the following step the weather and traffic data was combined. The functioning of the model is measured by Root Mean Square Deviation (RMSE) and lowers the value results in higher accuracy. The superior support vector machine model and the envisioned deep neural network model are correlated and the Deep neural networks has the higher accuracy. However, more research is expanded only for traffic prefigure rather than the incorporation of weather data.

3. Prediction Methods

In general traffic prediction methods are categorized into three types:-

1) Traffic model-based methods
2) Statistical methods
3) Machine learning-based methods

3.1 Traffic Model-based Methods

- Macroscopic traffic model
- Microscopic traffic model

Advantages:

- well-established theoretical background and good interpretability,
- consider the effect of various factors (incidents, road works, traffic control measures, etc).

Disadvantages:

- many parameters and assumptions need to be defined in advance,
- parameter calibration is difficult,
- small errors of these parameters could lead to large prediction errors.

3.2 Statistical Methods

- Local linear regression
- ARIMA (Auto-regression Integrated Moving Average)

Advantages:

- comparatively basic structure,
- easy and practical to be used.

Disadvantages:

- requires a great deal of historical data,
- gives some restrictive assumptions,
- difficult to model nonlinear traffic flow.

3.3 Machine Learning-based Methods

- ANN (Artificial Neural Network)
- SVM (Support Vector Machine)

Advantages:

- gains knowledge from training data,
- do not need to predefined model structure,
- good robustness.

Disadvantages:

- requires a training procedure and a large number of training data.

4. Research Gap

According to the reviewed literature:

- Several methods have been researched at a deep level for predicting the traffic flow and still an efficient method to satisfy real world traffic problem have not been identified.
- The predictive ability of a single forecast model is limited.
- There is a clear need for using robust models to determine the routes of vehicles which provide transport services inside cities.
- It is still a challenging work to select proper model in actual applications.

Now the question is, how can the short term flow of traffic be predicted using Deep Learning approach??

5. Objectives

The overall aim of this research proposal is to develop model which is able to accurately predict short-term traffic flow on urban arterial roads. The specific research objectives to fulfil this scope are:

- Develop robust models to accurately predict traffic during both normal and abnormal traffic conditions on urban arterial roads;
- Develop traffic prediction model that can be easily implemented without laborious calibration and maintenance, and that have the quality of location transferability;
- To develop a short term traffic flow prediction model using Deep Learning that brings high accuracy to satisfy real world traffic problems; and
- To propose a model to measure traffic and weather conditions.

6. Methodology

1. Design Process
   - ✓ Robustness is an important consideration for a traffic prediction model. Most short-term traffic prediction models learn from past data to make their predictions. The normal condition is occurring during planned events such as road works and

abnormal condition occur during unplanned events such as incidents and accidents. A robust prediction model should be accurate during both the conditions.

- ✓ A short-term traffic model should be easily implementable at different locations. Transferability refers to the ability of a model to work well with the less data.
- ✓ At the first stage data will be collected by locating sensor cameras at different areas in Jaipur city. The next stage will be the data pre processing stage, which involves combining the traffic and weather data. For the next stage, the traffic will be ready to run through the model. Next we will build the model by applying Deep Neural Network.

The process flow diagram involves following steps:

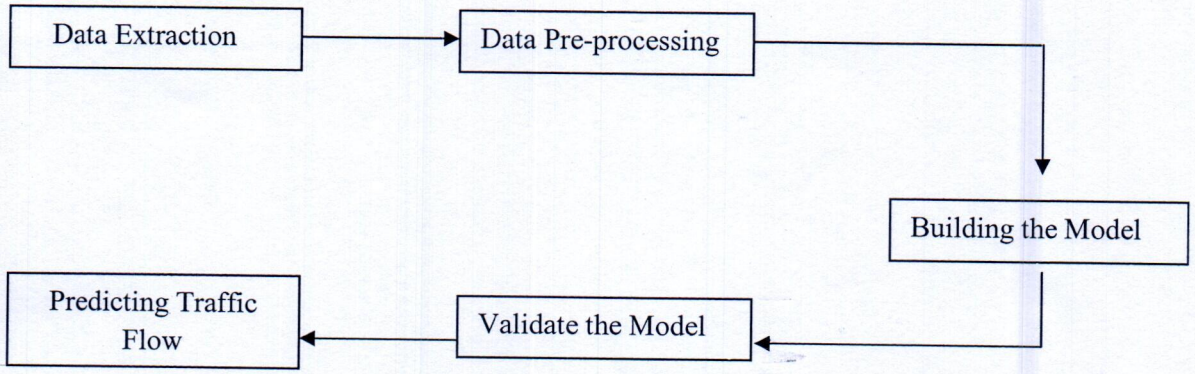

Fig 1: Process Flow Model

In summary, prediction models should be easily implementable and transferrable. Model implementation and transferability across locations will be investigated using traffic data collected from different road layouts under traffic signal controls and weather data.

2. Software

The software tools involved from the initial stage through to the final prediction are listed below:

- Google Refine - Initial data cleaning of traffic and weather data.
- Microsoft Excel – Initial data analysis and filtering of data.
- R studio – Predicting the data for future analysis.

3. Expected Outcomes

- The model will be robust and transferrable;
- High accuracy;
- Will be able to predict both traffic and weather conditions.

7. References

[1] Yuhan Jia, Jianping Wu, and Yiman Du, "Traffic speed prediction using deep learning method," IEEE 19th ITSC, Brazil, November 1-4, 2016.

[2] Han Jiang, Yajie Zou, Shen Zhang, Jinjun Tang, and Yinhai Wang, "Short-Term Speed Prediction Using Remote Microwave Sensor Data: Machine Learning versus Statistical Model," Hindawi Publishing Corporation Mathematical Problems in Engineering Volume 2016, Article ID 9236156.

[3] Daniel Kubek, Pawel Wiecek, Konrad Chwastek, "The impact of short term traffic forecasting on the effectiveness of vehicles routes planning in urban areas," ScienceDirect, Transportation Research Procedia 18(2016) 172-179.

[4] S. Vasantha Kumar, Lelitha Vanajakshi, "Short-term traffic flow prediction using seasonal ARIMA model with limited input data," Springer, Eur. Transp. Res. Rev. (2015) 7:21.

[5] Zhiheng Yu, Tieli Sun, Hongguang Sun, and Fengqin Yang, "Research on Combinational Forecast Models for the Traffic Flow," Hindawi Publishing Corporation, Mathematical Problems in Engineering Volume 2015, Article ID 201686.

[6] Sivabalaji Manoharan, "Short Term Traffic Prediction Using Deep Learning Approach," Msc Research Report in Data Analytics, 21st December 2016.

[7] S.Piramu Kailasam (Ph.d), K.Aruna Anushiya, Dr. M.Mohammed Sathik, "Traffic flow Prediction with Big Data Using SAE'S Algorithm," IJCSMC, Vol. 5, Issue. 7 July 2016, pg.186-193.

[8] Jia Zheng Zhu, Jin Xin Cao, Yuan Zhu, "Traffic volume forecasting based on radial basis function neural network with the consideration of traffic flows at the adjacent Intersections," Article in Transportation Research Part C Emerging Technologies · July 2014.

[9] Kartikeya Jha, Nishita Sinha , Shriniwas S. Arkatkar, and Ashoke K. Sarkar, "A comparative study on application of time series analysis for traffic forecasting in India: prospects and limitations," RESEARCH ARTICLES.

[10] Yiannis I. Kamarianakis and Poulicos P. Prastacos, "Space-Time Modelling Of Traffic Flow," 42nd Congress of the European Regional Science Association: "From Industry to Advanced Services - Perspectives of European Metropolitan Regions", August 27th - 31st, 2002, Dortmund, Germany.

Electronics and ICT Academy
(An Initiative of Ministry of Electronics and Information Technology, Govt. of India)

PDPM IIITDM JABALPUR

Faculty Development Programme on

Machine Learning Basics and Emerging Trends

September 01-06, 2017

This is to certify that **Ms. Pregya Poonia** from Mody University, Lakshamangarh, Sikar has attended and successfully completed the Faculty Development Programme on **Machine Learning Basics and Emerging Trends** during September 1-6, 2017 conducted by School of Computer Science and IT, Devi Ahilya Vishwavidyalaya, Indore and Electronics and ICT Academy, PDPM Indian Institute of Information Technology, Design and Manufacturing Jabalpur, and secured B grade.

Prof. Aparajita Ojha
Coordinator
E& ICT Academy, PDPM
IIITDM Jabalpur

Dr. Maya Ingle
Coordinator
School of Computer Science & IT,
DAVV, Indore